

CLAIMS

1. A method of identifying events in a process, the method comprising:
running a principle component analysis model on sensor data from the process;
calculating statistics related to the model;
determining if an event is occurring; and
finding a nearest cluster of bad actors related to the event to identify the event.

2. The method of claim 1 wherein finding a nearest cluster of bad actors comprises comparing the bad actor vectors to known clusters in a library of clusters for bad actors.

3. The method of claim 1 and further comprising for new bad actors:
identifying a sequence of cluster matches; and
correlating the sequence of cluster matches to known events.

4. The method of claim 3 and further comprising:
determining if a cluster needs to be split when new bad actors are added; and
splitting the cluster into two clusters using a goodness of fit algorithm.

5. The method of claim 4 and further comprising:
determining if a new event category is encountered; and
broadening limits for the sequence of clusters.

6. The method of claim 1 wherein a cluster is limited to a predetermined number of bad actors.

7. The method of claim 6, wherein the predetermined number of bad actors is ten.

8. The method of claim 1 wherein the statistics comprise Q and T2.
9. The method of claim 1 and further comprising using a feature scoring scheme to identify top contributors of bad actors.
10. The method of claim 9 wherein the feature scoring scheme is based on rank, value, and percent of contribution to a Q-residual sensor to identify a relative importance.
11. The method of claim 10, wherein the top-contributors are determined based on a majority percentage of the Q-residual.
12. The method of claim 10, where the top-contributors are determined based on only the contributors with absolute values that are drastically different from values of other contributors.
13. The method of claim 10 wherein the scoring scheme is based on predetermined limits.
14. The method of claim 13 wherein, the limits are computed statistically through change point detections.
15. The method of claim 9, wherein a predetermined minimum/maximum number of contributors are selected from rank, value, and percent of contribution to a Q-residual sensor to identify a relative importance.
16. A system for identifying events in a process, the system comprising:
means for running a principle component analysis model on sensor data from the process;

means for calculating statistics related to the model;
means for determining if an event is occurring; and
means for finding a nearest cluster of bad actors related to the event to identify the event.

17. The system of claim 16 wherein the means for finding a nearest cluster of bad actors comprises means for comparing the bad actor vectors to known clusters in a library of clusters for new bad actors.

18. The system of claim 16 and further comprising: (for new bad actors)
means for identifying a sequence of cluster matches; and
means for correlating the sequence of cluster matches to known events.

19. The system of claim 18 and further comprising:
means for determining if a cluster needs to be split (when new bad actor(s) are added); and
means for splitting the cluster into two clusters using a goodness of fit algorithm.

20. The system of claim 19 and further comprising:
means for determining if a new event category is encountered; and
means for broadening limits for the sequence of clusters.

21. The system of claim 16 wherein the statistics comprise Q and T2.

22. The system of claim 16 and further comprising means for feature scoring to identify top contributors of bad actors in a cluster.

23. The system of claim 22 wherein the means for feature scoring is based on rank, value, and percent of contribution to a Q-residual sensor to identify a relative importance.
24. The system of claim 23, wherein the top-contributors are determined based on a majority percentage of the Q-residual.
25. The system of claim 23, where the top-contributors are determined based on only the contributors with absolute values that are drastically different from values of other contributors.
26. The system of claim 23 wherein the scoring scheme is based on predetermined limits.
27. The system of claim 26 wherein, the limits are computed statistically through change point detections.
28. The system of claim 22, wherein a predetermined minimum/maximum number of contributors are selected from rank, value, and percent of contribution to a Q-residual sensor to identify a relative importance.
29. A system for identifying events in a process, the system comprising:
 - a controller coupled to sensors monitoring a process;
 - a principle component analysis model receiving data from the sensors monitoring the process and reducing a number of variables associated with the data from the sensors, the model further comprising:
 - a training module that is run on historical data to create a pool of vectors with values for the variables, wherein the training module further creates clusters of bad actors from the values based on statistics and associates the clusters with known events; and

a run time module that receives incoming data from the sensors, calculates statistics, determines if events are occurring, and identifies clusters to identify events.